DOCUMENT RESUME

ED 297 058 UD 026 288

AUTHOR Lam, Tony C. M.
TITLE Testing, Opportunity Allocation, and Asian and Pacific Americans. The Proceedings of a Hearing Co-Sponsored by the National Commission on Testing and Public Policy and the National Association for Asian and Pacific American Education. (Honolulu, Hawaii, April 11, 1987). A Report.
INSTITUTION National Commission on Testing and Public Policy.
SPONS AGENCY California Univ., Berkeley. Graduate School of Education.; Ford Foundation, New York, N.Y.; William and Flora Hewlett Foundation, Palo Alto, Calif.
PUB DATE Apr 88
NOTE 26p.; The Proceedings of a Hearing Co-Sponsored by the National Commission on Testing and Public Policy and the National Association for Asian and Pacific American Education.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Achievement Tests; *Asian Americans; Educational Assessment; Educational Policy; *Educational Testing; Elementary Secondary Education; Hawaiians; Individual Testing; Limited English Speaking; *Norm Referenced Tests; Policy Formation; *Standardized Tests; Students; *Test Bias; Test Construction; Testing Problems; Test Norms; Test Reliability; Test Validity
IDENTIFIERS *Asian American Students; Hawaiian Studies; *Pacific Islanders

ABSTRACT

The National Commission on Testing and Public Policy is conducting a 3-year policy-oriented investigation of the role of tests, especially standardized, norm-referenced tests, in the allocation of educational, training, and employment opportunities in the United States today. This document reports on the first hearing, which focused on the concerns of Asian and Pacific Americans (APAs). Testimony on the following issues is summarized: (1) Cultural and Linguistic Sources of Bias and Unfairness in the Informal Assessment of Asian Americans; (2) The Testing of Prospective Teachers in Hawaii; (3) Test Scores as Indicators of Educational Problems; (4) The Assessment of the Education of Native Hawaiians; (5) Modification of Standardized Tests Administered to Special APA Populations; (6) Testing and the Identification of Appropriate Educational Services for Limited English Proficient APA Students; (7) Research, Development, and Policy Issues Pertaining to Tests, Testing, and APAs; (8) Testing and APAs in California; (9) Testing Research and Test Taker Rights; (10) Ethics of Educational Testing; (11) Test Scores and the Ethnic Minority Index; and (12) Testing and the Educational Status of Native Hawaiians. Other testing issues of concern to APAs, but not covered in the hearings, are the following: (1) out-of-level testing; (2) special norms and renorming; (3) non-comparability of English language assessment tests; (4) testing in a language other than English; (5) test conditions and test sophistication; and (6) APA representation in test debias procedures. A list of 21 references is included. (BJV)

ED297058

TESTING, OPPORTUNITY ALLOCATION, AND ASIAN AND PACIFIC AMERICANS

The Proceedings of a Hearing co-sponsored by the National Commission on Testing and Public Policy and the National Association for Asian and Pacific American Education

A Report Prepared by Tony C. M. Lam

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Linda Wing
Nat'l Com. on Testing & Pub. Policy

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

UD026288

8-15-88

TESTING, OPPORTUNITY ALLOCATION, AND ASIAN AND PACIFIC AMERICANS

The Proceedings of a Hearing
co-sponsored by the
National Commission on Testing and Public Policy
and the
National Association
for Asian and Pacific American Education

A Report Prepared by
Tony C. M. Lam

for the

National Commission on Testing and Public Policy
Graduate School of Education
University of California
Berkeley, California 94720
April 1988

Acknowledgements

The National Commission on Testing and Public Policy wishes to thank the National Association for Asian and Pacific American Education for co-sponsoring the April 11, 1987 hearing on Testing, Opportunity Allocation, and Asian and Pacific Americans. We are particularly indebted to the following association members: (1) Dr. Amy Agbayani, Assistant to the Vice President for Academic Affairs at the University of Hawaii at Manoa, and Dr. Morris K. Lai, Educational Associate with the Curriculum Research Development Group at the University of Hawaii at Manoa, for their work in organizing the hearing; and (2) Dr. Tony C. M. Lam, Assistant Professor, Department of Educational Foundations, University of New Mexico, for his preparation of this report.

The preparation of this paper was supported by funds from the Ford Foundation, the William and Flora Hewlett Foundation, and the Graduate School of Education at the University of California at Berkeley. The analyses, conclusions, and opinions expressed herein do not necessarily reflect the positions, policies, or views of these organizations.

Introduction

The National Commission on Testing and Public Policy is a new, blue-ribbon panel supported by the Ford Foundation. It is conducting a three-year, policy-oriented investigation of the role of tests, especially standardized, norm-referenced tests, in the allocation of educational, training and employment opportunities in the United States today. As part of its study, the Commission is sponsoring a series of hearings to provide opportunities for various "testing publics" to present their findings, recommendations, and views about the impact of testing on American life.

The first hearing, held in Honolulu, Hawaii on April 11, 1987, and co-sponsored by the National Association for Asian and Pacific American Education, as part of its ninth annual meeting, focused on the concerns of Asian and Pacific Americans (APAs). APAs are the fastest growing population group in the United States (Gardner, Robey and Smith 1985). Specifically, the APA population grew by 142 percent between 1970 and 1980, and by another 48.5 percent between 1980 and 1985. Estimates indicate that APAs numbered 5.1 million in 1985, and projections show that the total will grow to 9.9 million by the year 2000.

In the period 1980-84, the highest proportion of immigrants who entered the United States through documented channels, some 48 percent, came from Asian countries. Approximately 59 percent of all APAs are foreign-born, with the figure rising to more than 70 percent for subpopulations such as Asian Indians, Koreans, and Vietnamese.

About two-thirds of the APA population speak an Asian language at home. Of this group, one-fifth of those aged five and above are thought to be limited English proficient (LEP).

On tests of verbal ability, the average score of APAs who are LEP is substantially lower than the average score of white students. It is likely that the limited English proficiency of these APA students depresses their scores on tests of knowledge, skills, and abilities in subject areas such as science and social studies. The linguistic and sociocultural characteristics of APA LEP students may be misinterpreted by non-APAs in assessments of APA performance in interviews and on application letters and essays.

There is no reason to believe that adult APAs who are LEP fare any differently on tests and assessments than APA LEP students. Moreover, there is evidence that many APA students and adults who are LEP are not receiving educational, training, and employment services commensurate with their linguistic and

sociocultural needs (see, for example, data cited in Tsang and Wing 1985).

The April 11, 1987 hearing was an opportunity for APAs to discuss their concerns about testing and its influence on the educational, training, and employment opportunities of their linguistically and socioculturally diverse community. This report summarizes the proceedings of the hearing. It also provides commentary on selected testing issues of concern to APAs that were not covered at the hearing.

Part I: Summary of the Proceedings of the Hearing

Twelve testimonies were given at the two-hour-long hearing. The presenters of the testimonies were:

- Dr. Chui Lim Tsang, Executive Director of the Chinatown Resources Development Center, San Francisco, California.
- Dr. Sheila Forman, Coordinator, Hawaii Mental Health Association.
- Selvin Chin-Chance, Testing Specialist, Hawaii State Department of Education, Honolulu, Hawaii.
- Dr. Ormond W. Hammond, Director, Office of Program Evaluation and Planning, Kamehameha Schools/Bernice Pauahi Bishop Estate, Honolulu, Hawaii.
- Dr. Tony C. M. Lam, Assistant Professor, Department of Educational Foundations, University of New Mexico.
- Dr. Pio DeCano, Counselor, Tacoma Public Schools, Tacoma, Washington.
- Rona Rodenhurst, Education Division Officer, Hawaii State Office of Hawaiian Affairs, Honolulu, Hawaii.
- Dr. Rosita Galang, President, National Association for Asian and Pacific American Education, San Francisco, California.
- Dr. Edmund Lee, President, California Association for Asian and Pacific Bilingual Education, Los Angeles, California.
- Dr. John Lum, Divisional Assistant, General Services, San Francisco Unified School District, San Francisco, California.
- Robert Hall, Hawaii Institute for Biosocial Research, Honolulu, Hawaii.
- Dr. Thomas T. Saka, Professor, Department of Educational Psychology, University of Hawaii at Manoa, Honolulu, Hawaii.

All of the testimonies were prescheduled. The presentation by Chui Lim Tsang was an invited address sponsored by the Commission.

Dr. Bernard R. Gifford, Dean of the Graduate School of Education at the University of California at Berkeley, and Chair of the National Commission on Testing and Public Policy, opened the hearing with an

overview of the objectives and activities of the Commission. Next, Chui Lim Tsang presented a paper entitled "Informal Assessment of Asian Americans: A Cultural and Linguistic Mismatch." Afterwards, other testimonies were given. Due to time constraints, these testimonies were limited to eleven in number and each lasted a maximum of ten minutes each. Forty-two individuals were present in the audience. Summaries of the paper and testimonies follow.

Cultural and Linguistic Sources of Bias and Unfairness in the Informal Assessment of Asian Americans.

In his invited paper, Chui Lim Tsang discusses the cultural and linguistic sources of bias and unfairness in the informal assessment of Asian Americans. Informal assessment is defined as the use of subjectively scored interviews and application letters and essays to select Asian Americans for educational and employment opportunities. Tsang confines his analysis to Asian Americans who are citizens or permanent residents of Asian descent and who come from non-English language backgrounds. Individuals with non-English language backgrounds include: 1) those who are limited in English speaking, reading, and writing skills; and 2) those who possess functional English skills in speaking and reading, but who use discourse strategies and conventions in formal writing or speech that are culturally and linguistically dissimilar to those of native English speakers. Tsang maintains that both of these types of linguistic characteristics are considered to be indicative of low ability when Asian Americans are considered for college entrance and job openings. Thus, cultural and linguistic bias operates to limit unfairly the educational and employment opportunities of Asian Americans.

Tsang provides illustrations of how Asian Americans who use speech and writing conventions that are derived from their cultural and linguistic backgrounds and experiences can be unfairly assessed in interviews and application letters and essays. For example, although Asian Americans may write letters and essays in grammatically correct English, they may write with a style, or with an organization, that is downgraded by native English speakers. Whereas individuals schooled in the United States are taught to state their topic sentence, or main thesis, at the start of their essays and letters, individuals schooled in Chinese countries are taught to do the opposite. That is, they marshall their facts and points to build an argument that culminates with the main point. This culturally and linguistically different discourse strategy is often considered "beating around the bush" and has been shown to be the basis for negative evaluations of the qualifications of Asian Americans for jobs and college admissions.

To support his contentions, Tsang cites findings from the research literature in linguistics and from real-world situations in

occupational licensing, job interviews, and college admissions essays. He gives particular attention to the potentially biased assessment of essays submitted by Asian American applicants to the University of California at Berkeley. The full text of Tsang's paper is included in a forthcoming volume of papers to be published by the National Commission on Testing and Public Policy.

The Testing of Prospective Teachers in Hawaii.

Sheila Forman criticizes the Hawaii state policy requiring the testing of prospective teachers. Forman maintains that the National Teacher Examinations are inappropriately or inaccurately designed to measure effective teaching in multilingual, multiethnic societies. In her view, the tests do not even consider the state of the art in the profession of teaching. The test scores cannot and do not tell us about the skills, abilities and competencies of teacher applicants in Hawaii.

Since a large number of the individuals who fail the test are the very ones urgently needed to teach in a linguistically and culturally diverse society, Forman criticizes the Hawaii State Department of Education (DOE) for not appropriating funds to provide help to those who are failing. In her view, the DOE is using teacher competency testing punitively to disqualify ethnic minority teacher applicants, including Filipinos, who are much needed in Hawaii.

Test Scores as Indicators of Educational Problems.

Selvin Chin-Chance reminds us that psychometricians and other testing experts are sometimes perceived by the public as "villains" because they often convey, via test scores, unwelcome news about the condition of education. He suggests that state departments of education have an obligation not only to collect and publicize test score data, but also to take steps to remedy the educational problems underlying low test score performance. In the case of high failure rates of Asian and Pacific American minorities on teacher competency tests, for example, Chin-Chance believes that the Hawaii State Department of Education should provide programs to assist individuals to acquire the knowledge and skills needed to pass the examinations.

The Assessment of the Education of Native Hawaiians.

Ormond Hammond presents three concerns about the educational assessment of native Hawaiian students. First, he argues for the end of the practice of subsuming native Hawaiians within the broad category of "Asian and Pacific American." Such data aggregation masks the unique problems of native Hawaiians. In particular, since native Hawaiian students are 21.7 percent of the public school population in Hawaii, the state should not

merge data on Hawaiians with data on other Asian and Pacific Americans in educational surveys, research studies, or program evaluations.

Second, Hammond notes that the health status of Hawaiian families places their children at educational "risk." The 1983 Native Hawaiian Educational Assessment Project, mandated by the United States Congress, found that Hawaiian students score below national norms on the achievement test series administered by the state of Hawaii on all subtests and at all grade levels. In the examination of the factors related to poor test performance and low educational achievement, the assessment project found that the high incidence of health problems among Hawaiian families may negatively affect their children's educational performance. For example, a large percentage of Hawaiian newborns are low birthweight babies; and the health problems that accompany low birthweight may affect their readiness for schooling. This is evidenced by the fact that many Hawaiian children who enter the preschools of the Kamehameha Schools in Honolulu fail the hearing screening test.

Third, Hammond emphasizes that scores on standardized norm-referenced tests have limited usefulness for the assessment of educational progress. Although standardized test score data provide information of value to general educational policy planning, Hammond says that "it should not be taken to mean that our job will be done when and if the Hawaiian 'curves' exactly match those of any given norm or reference group." As he sees it, a better approach to assessing educational outcomes over the long term is to employ "course-related tests" that match the curriculum and that are designed to measure achievement of specific instructional objectives.

Modification of Standardized Tests Administered to Special Asian and Pacific American Populations.

According to Tony C.M. Lam, when a standardized test is administered to Asian and Pacific American (A.PA) students whose linguistic, cultural and educational backgrounds are significantly different from those of the standardization group, the test scores are less reliable and valid than those reported by the test publisher. In order to restore this loss in psychometric quality, Lam recommends clarifying test instructions, extending test-taking time limits, and simplifying the language of test items. Although these modifications can enhance the reliability and validity of tests administered to special populations of APA students, the validity of the test norms is reduced at the same time. Research is needed to investigate the effects of modifying standardized tests as a way of increasing the

appropriateness of these tests for APA populations who are limited English proficient, who are newcomers to the United States, and/or who are unfamiliar with formal schooling.

Testing and the Identification of Appropriate Educational Services for Limited-English-Proficient Asian and Pacific American Students.

Pio DeCano believes that school procedures for placing limited-English-proficient (LEP) Asian and Pacific American (APA) students in appropriate instructional programs are inadequate. In his experience, personnel in most school districts use student scores on an oral English language proficiency test or a state-approved standardized achievement test to make decisions to enter or exit APA LEP students from bilingual and other special educational programs. In small school districts, the needs of APA LEP students are often assessed by educational personnel through conversations, or attempted conversations, with the students when they are enrolled in school. In DeCano's view, neither the administration of standardized English tests nor assessments based upon informal conversations conducted in English are capable of leading to principled decisions to place APA LEP students in instructional programs that meet their needs. In order to accomplish that goal, DeCano recommends assessing APA LEP students in their native languages.

DeCano considers "absurd" the administration of state- or school district-mandated standardized achievement tests, in subjects other than English, to APA LEP students. Because of the language barrier, these students cannot understand the test instructions or the test items. Yet they may possess a high level of knowledge and skills in the subject matter being tested. DeCano proposes that school districts consider exempting LEP students from state- and school district-required standardized achievement tests.

According to DeCano, it is no easy task to determine if an APA LEP student is experiencing educational problems due to the language barrier or due to a learning disability. He recommends observing and evaluating the student's progress during one complete school year before initiating the process of assessment for placement in a special education program. DeCano also emphasizes that a speaker of the student's native language should work with school district personnel in the special education assessment process.

Finally, DeCano laments the lack of adequate instruments for the assessment of gifted APA LEP students. However, he believes that the performance subtests of the Wechsler Intelligence Scale for Children, especially if they are administered in the students' native languages, might be of some use in helping to determine giftedness among LEP students.

Testing and the Educational Status of Native Hawaiians.

Representing the Office of Hawaiian Affairs, an agency created in 1978 by an amendment to the Hawaii State Constitution and responsible for coordinating federal and state programs involving native Hawaiians, Rona Rodenhurst points out that native Hawaiian students rank lower academically than any of the major ethnic groups in Hawaii. Although their test scores on standardized achievement tests have been increasing slightly each year, native Hawaiian students still fall below national norms. She links the performance of native Hawaiians on standardized achievement tests to the low socioeconomic status and poor health of the native Hawaiian community.

Rodenhurst observes that the formats, procedures, and contents of standardized tests do not match the ways that native Hawaiian students learn or the ways they exhibit their learning. Hence, she questions the appropriateness of standardized tests for this population. She comments, "Traditionally, tests have been used against us and [yet] we see a trend towards moving to a sensitivity, not only of evaluating the individual but also assessing the system that is requiring the evaluation mechanism. We urge you to keep in mind the various traditional ways of learning, many of which are not part of the western culture."

Research, Development, and Policy Issues Pertaining to Tests, Testing, and Asian and Pacific Americans.

Rosita Galang observes that there has been little research on testing and Asian and Pacific Americans (APAs). She argues that qualitative and quantitative analyses of linguistic, sociocultural, and other sources of differential test performance between APA students and other populations of students should be undertaken. APAs should not only be the subject of future testing research, but their expertise should also be called upon in the conduct of testing research on all population groups.

Galang states that the limited body of research on testing and APAs is matched by the limited availability of assessment instruments that have been specifically designed for APA students. The few that are available have not been evaluated systematically or widely disseminated.

Finally, Galang emphasizes that English language proficiency is one of the most important limitations on the performance of limited-English-proficient (LEP) students on standardized tests of achievement in subject areas other than

English. If the language proficiency factor is not taken into consideration, Galang warns, standardized tests can be used inappropriately with respect to APA LEP students and their scores on standardized tests can be misinterpreted. On behalf of the National Association of Asian and Pacific American Education, Galang forwards the concern that the misuse of tests and the misinterpretation of test scores are having adverse effects on the allocation of educational opportunities to APA students.

Testing and Asian and Pacific Americans in California.

Speaking on behalf of the California Association for Asian and Pacific Bilingual Education, Edmund Lee raises three issues. First, he questions the state policy requiring individuals to pass the California Basic Educational Skills Test (CBEST) prior to being certificated as teachers. He asks if the CBEST is capable of ascertaining the reading and writing abilities of Asian and Pacific Americans (APAs). He notes that many APAs pass the mathematics portion of the CBEST, but fail the English reading and writing portions. Lee states that 45 percent of the APAs who took the CBEST during the 1984-85 school year failed. Pointing out that there is a high demand for APA teachers in California—the current demand exceeds the supply by 56 percent—Lee maintains that the potential bias of the CBEST against APAs deserves close scrutiny. The California Association for Asian and Pacific Bilingual Education believes that the CBEST effectively serves to screen out many potential APA bilingual teachers who might otherwise be credentialed to teach in the state of California.

Second, Lee notes that no guidelines exist regarding the administration of standardized tests to limited-English-proficient (LEP) students. There is a pressing need for guidelines on when to include or exempt LEP student from standardized testing. This concern is prominent in California because approximately 130,000 APA students in the state are limited in English proficiency.

Third, Lee recommends the development of policies and procedures for the identification of APA students who may need special education. Currently, he states, there is no sound and standard policy for distinguishing APA students with exceptional needs from those who are simply in the process of acquiring English.

Testing Research and Test Taker Rights.

John Lum recommends that test producers and test users make it possible to identify test takers by race and ethnicity. Without data on Asian and Pacific American (APA) test takers by ethnicity (Chinese, Japanese, Filipino, Korean, Vietnamese, and so on), Lum points out, it is impossible to conduct item-analyses and other types of studies that might lead to a better understanding of the factors associated with the test performance of specific groups of APA students and of how their test performance might be enhanced.

Additionally, Lum advocates the adoption of some form of expanded scale scores, that is, the placement of scores obtained from different achievement tests on one common scale. Such a practice would enable longitudinal analyses of the educational progress of students from the primary grades through high school graduation.

Lum calls for routine disclosure of test scores to students and their parents. He describes a plan whereby school districts would provide families with forms upon which their children's scores on standardized tests, from first grade through the twelfth grade, would be regularly recorded. This policy would enable students and their parents to gain control over the use and interpretation of the students' test scores. For example, the students and their parents could ask school personnel to verify the accuracy of the scores and to reveal what decisions are made based upon the scores.

Ethics of Educational Testing.

Robert Hall argues against the continued use of standardized tests for college admissions. He maintains that educational aptitude tests have little or no predictive validity. Additionally, since test companies use the test results for research, test takers, in Hall's view, are being studied without their consent or knowledge. Hall calls for a multidisciplinary examination of the moral, ethical, and legal issues involved in testing.

Test Scores and the Ethnic Minority Index.

Thomas Saka addresses the issue of the "ethnic minority index" used in comparing standardized test scores across school systems. He notes that, when student scores on aptitude and achievement tests are used as one basis of comparing the quality of education across states, the scores are often adjusted by the percentage of black and Hispanic students in a state.

Saka does not advocate the interstate comparison of test scores to assess educational systems, even when adjustments are made for the percentage of ethnic minorities in the student population. In his view, the causes of differential test performance are not consistent across racial and ethnic groups. If, however, such interstate comparisons are made, then Saka recommends that the ethnic minority index include the percentage of Asian and Pacific American (APA) students in the state population, along with the percentages of blacks and Hispanics. States with high percentages of APA students, such as Hawaii, would then have markedly different results than when minority indices including only blacks and Hispanics are used.

Summary of Concerns and Recommendations that Emerged from the Hearing.

Many of the issues presented in the testimonies are directly or indirectly related. The concerns and recommendations that were expressed are grouped into twelve main issues and highlighted below.

(1) The use of culturally and linguistically different discourse conventions by Asian and Pacific Americans (APAs) in interviews, letters, and essays required for college admission or employment can be negatively evaluated by non-APAs. To the extent that these cultural and linguistic differences do not relate to criterion performance, APA access to equal educational and economic opportunities is unfairly limited.

(2) APAs are failing teacher competency tests at high rates. These tests should be scrutinized to determine if they measure the knowledge and skills that competent teachers must possess. They should also be examined for potential bias against APA teacher applicants.

(3) The low performance, on average, of native Hawaiian students on standardized achievement tests may be related to the health problems of native Hawaiian families. It is also the case that native Hawaiian students, due to their cultural background, learn and exhibit learning in ways that are not measured by standardized norm-referenced tests.

(4) Specially designed tests that measure what is taught in the curriculum are more useful for improving educational programs at the local level than are standardized tests that are commercially available or that are designed for a national assessment. On the other hand, if the scores of all commercially available standardized tests used by a school district were placed on a common scale, the data might be

useful in studying the educational progress of students over the course of their entire elementary and secondary school education.

(5) Test score data, indeed all educational data, should be collected and disaggregated by race and ethnicity. Each racial and ethnic group has distinct characteristics and needs. Accordingly, APAs should not be subsumed into the broad category of "other;" and native Hawaiians should not be subsumed into the broad category of "Asian and Pacific Americans."

(6) To increase the validity of standardized tests administered to APA students who are limited English proficient, slight modifications of the testing procedures and test items should be made.

(7) APA students who are limited in English proficiency are often not placed into educational programs that address their talents and needs due to the use of inappropriate assessment procedures and the misinterpretation of test scores. Appropriate procedures include the use of the student's native language to conduct the assessment. Linguistic and sociocultural background differences between the APA population and the standardization group should be considered when using APA test scores to make educational decisions.

(8) School districts and other educational agencies should establish common guidelines for the exemption of limited-English-proficient students from standardized testing in subject areas other than English.

(9) Assessment instruments specifically designed for APAs need to be developed.

(10) Research to study the determinants of the test performance of APAs, especially those who are limited English proficient, is much needed. APA researchers should be involved in the study of testing and public policy issues as they relate to all population groups.

(11) The rights of test takers should be recognized by school districts that use tests and by manufacturers of tests. Students and their parents should routinely be provided with the students' test scores so that they can use the information to participate in and influence educational decisions affecting the students' schooling. Test takers should have the option of barring test manufacturers from using their test scores for research purposes.

(12) When comparing standardized test scores across school systems or states, adjustments to test scores should be based not only upon the percentages of black and Hispanic students in a state, but also the percentage of APAs.

One critical issue that runs through most of the testimonies deserves additional discussion. Nearly all of the testing issues of concern to APAs are related to differences in the linguistic and sociocultural backgrounds of APAs and the majority population. The degree of difference is directly proportional to the extent of test performance difficulties experienced by APAs. It is also directly proportional to the extent to which the performance of APAs on tests, interviews, and application letters and essays may be misevaluated.

Linguistic characteristics include both levels of English proficiency and familiarity with the discourse conventions and strategies of native English speakers. Sociocultural background factors refer to the country of nativity, ethnicity, as well as the kind of family, educational, and community environments, beliefs, and values that shaped the student. When discussing testing policy and conducting testing research that will affect APA access to educational, training and employment opportunities, the specific subpopulation of APAs, by linguistic and sociocultural background, must be defined. Generalization of testing policies and testing research findings from one subgroup of APAs to another with a differing configuration of linguistic and sociocultural characteristics should be made with extreme caution.

It should also be emphasized that, while biased assessments of APAs may be inadvertent, due to lack of knowledge and understanding of their linguistic and sociocultural characteristics, the effect of inadvertent bias is the same as deliberate bias. APA access to equal educational and economic opportunities is unfairly limited.

Part II: Other Testing Issues of Concern to Asian and Pacific Americans

In this part of the report, selected issues related to the testing of Asian and Pacific Americans (APAs) that were not touched upon during the necessarily brief public hearing are presented. These issues are primarily related to APAs who are limited in English proficiency and who are unfamiliar with the American culture and mores. They concern: 1) out-of-level testing; 2) special norms and renorming; 3) the non-comparability of tests of English proficiency; 4) testing in a language other English; 5) test conditions and test sophistication; and 6) APA representation in test debias procedures.

Out-of-Level Testing.

In standardized achievement testing, examinees are assigned to test levels based on their enrollment in a particular grade level in school. This practice assumes that both the curriculum and the achievement levels of the examinees are typical for that grade level. However, the linguistic and sociocultural backgrounds of Asian and Pacific American (APA) students who are recent immigrants to the United States may differ so much from those of mainstream students that, for these APA students, on-level testing may not be appropriate in either content or difficulty. The English language proficiency of these APA students is likely to be much lower than that of other students in the same grade level. Also, the schooling these APA students obtained prior to coming to the United States is probably quite different from the schooling provided in this country. Some APA immigrant students may not have had any formal schooling in their native country. As a consequence, test scores obtained from on-level testing of APA immigrant students who are limited English proficient (LEP) may be largely a function of guessing. As such, the test scores are unreliable and invalid.

Apart from exempting APA LEP immigrant students from testing altogether, a potential solution is to administer achievement test levels that match these students' functional levels rather than their grade levels. For example, a fifth grade student reading at the fourth grade level would be given a fourth grade reading test rather than a fifth grade reading test. However, the student's test performance would be compared to the fifth grade norm group instead of the fourth grade norm group.

Out-of-level testing is currently a controversial issue. Proponents claim that measurement errors are reduced if test levels are assigned according to functional level, not grade level. Opponents criticize out-of-level testing on four grounds, namely that: 1) it is not a sound testing practice because of low

curricular and face validities; 2) equating different test levels by using an expanded standard score metric is suspect (see, for example, Slinde and Linn 1977);[1] 3) out-of-level testing is not necessary because even chance level scores can still predict grade point average (see, for example, Boldt 1968); and 4) testing students at lower levels may lower teacher expectations of student achievement.

Additionally, a problem that renders out-of-level testing especially problematic with respect to APA LEP immigrant students is the discrepancy in their quantitative and verbal achievement levels. Generally, APA LEP immigrant students acquire quantitative skills faster than English verbal skills. This phenomenon requires that their functional level in each skill area be determined. However, test publishers typically prescribe only one procedure for determining a single functional level for all subtests. Usually, this procedure involves selecting a functional level based upon a student's score on a short cognitive test consisting of mathematics and language arts items similar to those that appear in the full test battery.

Research is needed to determine the appropriateness and effectiveness of out-of-level testing for assessing the achievement of APA LEP immigrant students. Policy studies must include attention to the promise and problems of administering different test levels in mathematics and reading.

Special Norms and Renorming.

Opponents of the standardized testing of minority students often center their criticisms on the use of norms. Because of differences between the standardization group and the minority group, or the inadequate representation of minorities in the standardization group, norms furnished by test publishers are judged inadequate for determining the achievement status of minority students. A popular position is to recommend the development of special norms for minority groups or the renorming of existing norms to reflect the proportion of minority groups in the national population.

From one perspective, the establishment of special norms for Asian and Pacific American (APA) students is equivalent to the establishment of lower standards for APA students. To accept lower standards is to stigmatize APA students as having less potential to achieve than the standardization group. Special norms and lower

[1] An expanded standard score metric is a scale specifically developed to link or equate successive levels of a test (e.g., the CTBS Reading Comprehension Test), so that comparisons can be made between test scores obtained from different levels of the test. By using an expanded standard score metric, an examinee would obtain the same standard score regardless of which level of the test is administered.

standards become the "basis for invidious comparisons between racial groups" (DeAvila and Havassy 1974). According to Clark (1969), advocating the development of special norms for minority students is equivalent to saying "that these children are doomed, expendable in terms of not being capable of assimilation into the economic, educational, and cultural norms of the larger society." Stigmatization is hardly the end result envisioned by the advocates of special norms. However, many contend that stigmatization would inevitably accompany the establishment of separate norms for minority groups.

On the other hand, it might be useful to use special norms in conjunction with regular norms. The special norms could provide a means of making finer distinctions among the performances of students within the same minority group. Comparing the scores of APA students to special norms and the regular norm could provide a sense of their academic standing relative to their own subpopulation group and to the national population.

However, the number of norms needed for APA students would render the development of special norms difficult. Norms would have to be developed for each of the APA subgroups formed by the combination of English language proficiency levels and sociocultural characteristics, especially ethnicity. The number of norm tables required could be so large that the effort would be beyond practicality.

The issue of inadequate norming for APAs can be viewed from the perspective of inadequate standardization. In test standardization, the psychometric properties of the test are based upon the test scores obtained from a standardization group. If the quality of the test is judged acceptable, norms are then developed based upon the performance of another standardization, or norm, group. The extent to which the APA population and the standardization groups differ on key characteristics and the extent to which APAs are represented in the standardization groups will determine the reliability and validity of the test for APAs.

Jensen (1980) recommends standardization within subgroups, that is, performing comparable item selection procedures separately within each subgroup, and combining the subgroups in proportion to their numbers in the general population in the final norming, when normalized standardized scores are computed. This standardization procedure is not designed to equalize the mean scores of subgroups "but at achieving highly similar reliability, factorial composition, predictive and construct validity, and range of item difficulties while at the same time minimizing item X subgroup interactions (i.e., the items should have the same rank order of difficulties across subgroups)."

As long as performance on norm-referenced tests is used to determine APA eligibility for educational and economic opportunities, the issues of special norms and renorming should be subjected to policy debate and systematic analysis. Both the psychometric and political implications of special norms and renorming must be investigated.

Non-Comparability of English Language Assessment Tests.

In 1974, the United States Supreme Court, in *Lau* v. *Nichols*, held school districts responsible, under Title VI of the 1964 Civil Rights Act, for providing special educational services to students who are limited in English proficiency due to their non-English language backgrounds. In order to provide the mandated special educational services, school districts must first identify and assess the English language proficiency of limited English proficient (LEP) students.

Many language assessment instruments have been developed. Their proliferation has led to studies to determine the comparability of a few of the most popular commercially available tests (Gillmore and Dickerson 1970, Ulibarri *et al.* 1980, Pelavin Associates 1986).

The results from these studies consistently suggest poor comparability among the instruments investigated. Different tests classify different percentages of students into LEP and English proficient (EP) categories, mainly because the linguistic abilities tapped by these tests are not comparable (Wald 1981). In the most recent study by Pelavin Associates (1986), two commonly used language proficiency tests, the Language Assessment Scale (LAS) and the Language Assessment Battery (LAB), were compared. Of the 910 students in the study sample who spoke some English, over 80 percent of these students were classified as LEP by the LAS, but as EP by the LAB.

The non-comparability of language proficiency tests has two ramifications. First, performance on language proficiency tests is generally used as the basis for decisions to enter or exit students into and from English as a second language (ESL) or bilingual education programs. When different tests are used by different districts, the entry and exit criteria vary as a result of test non-comparability. Given limited information on the predictive validity of these tests, districts that use a less stringent test such as the LAB may be denying educational services to students who need extra assistance to realize fully their potential to learn and succeed in school.

The second ramification of the non-comparability of language proficiency tests concerns decisions to exempt LEP students from school district- or state-mandated standardized achievement testing.

In some school districts and some states, including New York, students classified as LEP are exempt from having to take standardized achievement tests. Again, depending on the language proficiency test used, a student with a certain English proficiency level may be exempt in one district, but not in another. In this way, the validity of language proficiency tests will affect the validity of standardized achievement tests. If a less stringent language test is used to identify LEP students, a student who is not ready to take a standardized achievement test may, in fact, be tested. As a result, the score obtained on the achievement test will not be valid. Inappropriate educational decisions may be based on these invalid test scores.

The influence of the use of non-comparable English language proficiency tests on the educational opportunities of Asian and Pacific American LEP students warrants serious study. Both the direct influence of such tests on access to ESL and bilingual education programs and the indirect influence of such tests on achievement testing and achievement test score-based educational decision making need immediate attention.

Testing in a Language Other than English.

In the *Standards for Educational and Psychological Testing* (1985), it states that "for a non-native English speaker and for a speaker of some dialects of English, every test given in English becomes, in part, a language or literacy test . . . test results may not reflect accurately the abilities and competencies being measured if test performance depends on these test takers' knowledge of English." When a student does not have sufficient English language skills to take standardized tests in English (other than language proficiency tests), one possible way of determining the student's ability and achievement levels is to administer a test written in the student's native language. This test can be developed from scratch, or by translating an existing test. Although preferable, test construction requires enormous effort and vast resources that are usually unavailable to school districts; therefore, test translation is often undertaken. However, to produce an equivalent test in a non-English language requires more than mere translation and superficial adaptation. Some words and concepts that exist in English do not have equivalent counterparts in other languages. Direct translation may change the meaning and difficulty of test items (DeAvila 1973, Laosa 1973, Olmedo 1981, Samuda 1975, and Sechrest, Faye and Zaidi 1972).

Placement of students into educational programs should be based on the students' achievement profiles, their intellectual abilities, and other background information. For recent Asian and Pacific American (APA) immigrants who have no or very limited

English skills, it seems logical to assess their cognitive abilities and academic achievement with tests written in their native language, if they are fluent in their native language, and to place them into programs based on these test results, as well as other information. Moreover, if students are placed into programs in which one of the languages of instruction is their native language, then their progress should be assessed by tests written in their native language. However, there is a dearth of achievement tests written in Asian languages, and it is extraordinarily difficult to develop such tests. Short of exempting APA students who lack English proficiency from taking achievement tests, it is not clear what testing policies and procedures should be followed.

Test Conditions and Test Sophistication.

The conditions under which standardized tests and other forms of assessment are administered can negatively affect the performance of Asian and Pacific Americans (APAs) who are new to the United States and unfamiliar with the American school system. This bias due to test conditions may be reduced if the student is tested in a familiar setting by an examiner who speaks the student's native language and with whom the child has rapport.

A problem related to test conditions is the test sophistication of APA students. Research has shown that in addition to English language proficiency, test-taking experience is a contributor to test performance (see, for example, Millman, Bishop and Ebel 1965; Oakland 1972). APA students who are immigrants or refugees may not perform well on tests due to their lack of experience and skills in taking tests in the American educational system. These students may be "unfamiliar with the terminology, format, and procedures that are an integral part of standardized tests" (Blakely 1986). To eliminate bias due to lack of test sophistication, these students should be trained in test-taking skills.

In order that test scores reflect accurately the knowledge and abilities the standardized test is designed to measure, efforts must be made to avoid contamination by testing conditions and test-taking skills. Research studies and policy analyses should be conducted to examine test conditions and test-taking skills that would render standardized tests fair for APA immigrant and refugee students.

Asian and Pacific American Representation in Test Debias Procedures.

Almost all manufacturers of standardized achievement and ability tests ask panels of minorities to participate in test item review and tryout procedures leading to revisions designed to reduce test bias. To the extent that one minority group does not have a representative

on the review panel, bias against that group may not be averted. Efforts should be made to determine the level of participation of Asian and Pacific American (APA) reviewers in the test review panels of the major test publishers and to ensure that APA reviewers are adequately represented on such panels.

Conclusion

Standardized tests are not free from bias and misuse. However, it should be remembered that the phenomenon of group differences in test performance, in itself, does not constitute sufficient evidence of test bias. Careful and detailed studies are needed to identify the sources of observed differences in test performance between groups.

It is hoped that this report will contribute to the efforts of the National Commission on Testing and Public Policy to examine and make policy recommendations regarding the improvement of test making and the more informed use of tests in the allocation of educational, training and employment opportunities for Asian and Pacific Americans.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. 1985. *Standards for Educational and Psychological Testing.* Washington, D.C.: Author.

Blakely, M. M. February 1986. Teaching test-taking skills to elementary LEP students. *National Clearinghouse for Bilingual Education Forum.*

Boldt, R. F. 1968. Study of linearity and homoscedasticity of test scores in the chance range. *Educational and Psychological Measurement* 28:47-60.

Clark, K. B. November 1969. An interview with Dr. Kenneth B. Clark. *Council for Basic Education Bulletin* 14:8-18.

DeAvila, E. A. 1973. IQ and the minority child. *Journal of the Association of Mexican American Educators* 1:34-38.

DeAvila, E. A. and Havassy, B. November 1974. The testing of minority children–a neo-Piagetian approach. *Today's Education.* 63:72-75.

Gardner, R. W.; Robey, B. and Smith, P. C. October 1985. Asian Americans: Growth, Change, and Diversity. *Population Bulletin* 40:1–44.

Gillmore, G. and Dickerson, A. D. 1980. The relationship between instruments used for identifying children of limited English proficiency in Texas. *Bilingual Resources* 3:16-29.

Jensen, A. R. 1980. *Bias in Mental Testing.* New York: The Free Press.

Lau v. *Nichols,* 414 U.S. 563 (1974).

Laosa, L. M. 1973. Cross-cultural and sub-cultural research in psychology and education. *Inter-American Journal of Psychology* 7:241-248.

Millman, J.; Bishop, C. H. and Ebel, R. 1975. An analysis of test-wiseness. *Educational and Psychological Measurement* 25:707-725.

Oakland, T. 1972. The effects of test-wiseness materials on standardized test performance of preschool disadvantaged children. *Journal of School Psychology* 10:355-360.

Olmedo, E. 1981. Testing linguistic minorities. *American Psychologist* 36:1078-1085.

Pelavin, S. H. 1986. *Preliminary report on the Title VII project student entry and exit criteria study.* Paper presented at the annual conference of the American Educational Research Association, Washington, D.C.

Samuda, R. J. 1975. *Psychological Testing of American Minorities, Issues and Consequences.* New York: Harper and Row, Inc.

Sechrest, L.; Faye, T. L. and Zaidi, S. M. H. 1972. Problems of translation in cross-cultural research. *Journal of Cross-Cultural Psychology* 3:47-56.

Slinde, J. A. and Linn, R. L. 1977. Vertically equated tests: fact or phantom? *Journal of Educational Measurement* 14:23-32.

Tsang, S. L. and Wing, L. C. Winter 1985. *Beyond Angel Island: The Education of Asian Americans.* New York: Clearinghouse on Urban Education, Institute for Urban and Minority Education, Teachers College, Columbia University.

Ulibarri, D.; Spencer, M. L. and Rivas, G. A. 1980. *Comparability of three oral language proficiency instruments and their relationship to achievement variables.* San Francisco: American Behavioral Research Corp.

Wald, B. 1981. *On assessing the oral language ability of limited English proficient students. The linguistic bases of the non-comparability of different language proficiency assessment measures.* Paper presented at the Language Assessment Institute, National College of Education, Evanston, Illinois.